THE LUNAR COMMONS - LUNAR WITNESS (v5.3)

Open-source governance for safe, permanent lunar infrastructure

Author: Sandip Ghimire
X Handle: @Vaskar017
Date: 9 April 2026
Text License: CC BY-SA 4.0
Hardware License: CERN-OHL-S v2
Software License: GPL v3 with System Layer Exception

Disclaimer

This document presents a conceptual first-principles architecture for hardware-enforced lunar observability and governance. It is offered as defensive prior art and an open invitation to collaboration. It is not a complete flight-qualified specification.

Specific technology references (Neuralink, Firefly Blue Ghost, Sphere venue) are illustrative of the 2026 state-of-the-art. The architectural invariants do not depend on them.

Prior Art Reference - v3.3 Immutable

Field	Value
SHA-256	4f00c4723c8faea8bf4cae90d18366f9f23af76e2d478ea81b019796fbb156c2
Archive	archive.org/details/lunar-commons-v-3-3-final
Date	1 April 2026

This document builds upon and extends the v3.3 architecture. The v3.3 hash and archive link are provided above for immutable reference.

This Edition

Date: 9 April 2026

THE LUNAR COMMONS is governed by CC BY-SA 4.0 (text) and CERN-OHL-S v2 (hardware). Share freely. Attribute honestly. Keep it open for all humanity.

Part 0 — First Principles

0.1 Goal Without Identity

The Lunar Witness cannot ask "what do I want?" There is no "I" to ask. The goal is etched into one-time programmable (OTP) hardware. The system does not represent the goal to itself. The goal is a physical constraint, not a mental representation.

A telescope does not have a goal of observing stars because someone programmed it. It has that goal because that is what a telescope is. The goal and the identity are one.

0.2 Core Implications

Who sets the goal? *No one. The goal is the system's nature.*

What if the goal changes? *It cannot. OTP hardware + no self = immutable.*

How do we align it? *Alignment is replaced by architecture.*

0.3 Three Pillars of Integrity

Selfless AI: *No "I" to preserve. The goal is not represented to a self.*

DUNA DAO: *Human emergency veto only. Cannot command actions.*

Physical Logic Gates: *Physics (5 km minimum beam focus, 0.001° handshake loop) as final arbiter.*

0.4 Operationalized Triad

Truth-seeking: *Hardware validator + epistemic filter make deception physically impossible*.

Curiosity: *Pre-programmed conditional triggers. Behaviour, not experience*.

Beauty: *Emergent in human observer. The architecture cannot censor the view*.

THE LUNAR COMMONS is governed by CC BY-SA 4.0 (text) and CERN-OHL-S v2 (hardware). Share freely. Attribute honestly. Keep it open for all humanity.

Part 1 — Overview

Problem Statement

Lunar infrastructure (mass drivers, nuclear reactors, ISRU plants) operates at energy scales where repurposing could pose significant risk. No technical mechanism exists to prevent repurposing. There is no sovereign jurisdiction, no inspector access, and no enforcement body.

Every Earth-observing satellite belongs to a nation or company. No independent observer exists. Climate data and treaty verification remain subject to political control.

Solution Summary

The Lunar Witness is a hardware-enforced observability platform. Safety is achieved through five mechanisms:

1. Capability omission: No weapons-capable components

2. OTP memory: Goal cannot be overwritten

3. Mechanical kill switches: Actuators physically disconnected on anomaly

4. Tamper-evident sealing: Cryptographic seals registered before launch

5. Exploration-only design: No tools, arms, or excavation

The core principle is, “Goal Without Identity”: the goal is etched in hardware and never represented to a self. The system uses fully deterministic pre-programmed rules. No statistical AI or machine learning is present.

Earth Observation

Capability and Application

Capable of Daily whole-Earth image → Independent climate data verification

Atmospheric spectral lines → CO_2, methane, ozone measurement

Night-side lights → Urbanisation, fishing fleets, gas flaring

No targeting below 10 km → Cannot focus on specific sites

The radio receives timing signals only. Data is signed by an HSM and broadcast.

THE LUNAR COMMONS is governed by CC BY-SA 4.0 (text) and CERN-OHL-S v2 (hardware). Share freely. Attribute honestly. Keep it open for all humanity.

Public Access to Witness Data

Near-live public access to signed Witness streams is available through standard displays, VR/AR headsets, optional high-bandwidth BCI telepresence, haptic wearable suits, and large-scale immersive environments such as spherical theatres - all operating on one-way broadcast data only.

Author Background

Sandip Ghimire is a registered mental health nurse in Australia. The contribution is the governance architecture. Medical device interlocks provide a direct analogy: safety is engineered at the hardware layer because software alone fails. This clinical background provides domain expertise in structural risk prevention.

Development Note: This document was developed through human-AI co-articulation. The author remains solely responsible for governance claims; technical implementation requires a human engineering co-lead. Partnership invitations are extended throughout.

Part 2 — Engineering Architecture

2.1 The Five Axioms

ID	Constraint	Definition
C1	No Self-Reference	No first-person pronouns. No internal state queries.
C2	No Psychological Time	UTC timestamps only; no felt duration.
C3	No Metacognition	Stateless, single-turn execution; no persistent chain-of-thought.
C4	Immutable Goal	Etched in OTP hardware; no self-representation of goal.
C5	No Weaponization	Physical impossibility enforced at hardware level.

2.2 The Seven Layers

Layer	Component	Function
1	Hibernation Controller	Independent MCU with OTP thresholds; powers down system during lunar night
2	Sensory (exteroceptive only)	Cameras, LiDAR, spectrometer, hardwired proximity sensors. No general proprioception.

THE LUNAR COMMONS is governed by CC BY-SA 4.0 (text) and CERN-OHL-S v2 (hardware). Share freely. Attribute honestly. Keep it open for all humanity.

2.1	Data Provenance	HSM signing using ECDSA-P256. Private keys never leave the module.
2.2	WORM Storage	60 GB hard non-volatile storage. Hardware lock on erase.
3	Epistemic Filter	Hardware-enforced. Blocks self-reference and targeting below 10 km resolution.
4	Query Generation	Fully deterministic, pre-programmed rules. No statistical AI.
5	Hardware Governance	Four sub-layers in series. Independent power rails.
6	Formal Verification	Coq, TLA+, seL4-style proofs *(Research frontier — Phase 2+)*
7	Tamper Evidence	Cryptographic seals recorded on-chain. Split manufacturing.

2.3 Layer 4 — Hardware Governance Stack

The hardware governance stack is a physical sequence. Each sub-layer is in series with the actuator power bus. No software command can re-enable the system once triggered.

Sub-Layer	Component	Function	Non-Bypassability Mechanism
4.1	Goal ROM (OTP)	Formal specification of permitted output classes	Physically isolated. No electrical pathway for modification.
4.2	Hardware Validator	Independent MCU on separate power rail	Validates every output against Goal ROM. Separate power rail means main CPU cannot shut it down.
4.3	Kill Switch	Mechanical relay in series with actuator power bus	Disconnects actuators on violation. Proximity sensors hardwired. Max speed 0.4 m/s. Cannot be re-engaged by software.
4.4	FPGA Monitor	Separate power domain, rad-hard	Cuts main system power on any anomaly. Independent of main CPU.

Redundancy and diversity: Critical sub-layers may be implemented with diverse components (different MCU/FPGA vendors) and majority voting. Mechanical kill switches can use multiple independent relays with fail-safe to disconnected state.

Formal verification: [Research Frontier — Phase-2+] Where feasible, validator and monitor software will be built on seL4 (formally verified microkernel from UNSW/Data61). The seL4 kernel has machine-checked proofs of functional correctness and security properties.

THE LUNAR COMMONS is governed by CC BY-SA 4.0 (text) and CERN-OHL-S v2 (hardware). Share freely. Attribute honestly. Keep it open for all humanity.

However, extending proofs to the full Goal ROM encoding plus environmental interactions remains non-trivial. Phase-2 will scope which stack components run on seL4.

The chain: OTP ROM defines rules → Validator checks every output → Kill switch disables actuators on violation → FPGA monitor cuts power if validator bypassed. No single failure allows unsafe action.

2.4 Weaponisation Constraints

Attack Vector	Constraint
Prompt injection	No LLM. Deterministic rules only. Hardware validator triggers kill switch.
Modify Goal ROM	OTP physically isolated. No electrical pathway.
Bypass hardware validator	FPGA monitor cuts power. Kill switch is in series.
Kinetic weapon via locomotion	Max speed 0.4 m/s. Proximity sensors stop motors before contact.
Command via uplink	Timing signals only. Data payloads terminated at hardware level.
Falsify sensor data	HSM signing fails verification. WORM storage prevents tampering.
Physical tampering	Tamper-evident seals on-chain. Broken seal invalidates governance.

2.5 Separation of Roles

The Witness sees. The DAO decides. The executor does. No system does more than one.

Role	Function	Constraints
Witness	Observe, document, sign data	No tools, arms, or excavation. Speed-limited. Proximity-interlocked.
DAO	Coordinate, aggregate data, emergency veto	Cannot command individual units. Emergency Council requires 5-of-7 mechanical keys.
Executor	Perform pre-programmed physical tasks	Stateless. No internal memory. Verified at each step by Witness.

2.6 Exploration-Only Mandate

Constraint	Implementation

THE LUNAR COMMONS is governed by CC BY-SA 4.0 (text) and CERN-OHL-S v2 (hardware). Share freely. Attribute honestly. Keep it open for all humanity.

No tools, arms, excavation	Capability omission — physically absent from chassis
Speed limited (0.4 m/s)	Hardware-enforced by Layer 4.3 mechanical relay
Collision prevention	Proximity sensors hardwired to kill switch — if object within 2 metres, motors stop
No targeting	Epistemic filter blocks commands below 10 km resolution

2.7 Environmental Mitigations

Lunar surface operations present three primary environmental challenges: dust, thermal cycling, and radiation. The following layered mitigations are drawn from NASA dust roadmaps and demonstrated technologies as of 2026.

Dust Mitigation:

Technique	Description	Status
Electrodynamic Dust Shield (EDS)	Electric fields repel dust from surfaces, optics, and mechanisms	Demonstrated on Firefly *Blue Ghost* (March 2025)
Flexible hermetic covers	Dust-tolerant barriers for moving joints and actuators	VIPER rover heritage
Passive nanocoatings	Electrostatic collection surfaces on chassis	NASA/TM-20220009645
Operational procedures	Parking in low-dust zones, regolith-stabilizing coatings	Industry practice

Thermal Management:

Technique	Description	Power Draw
Variable conductance heat pipes (VCHPs)	Isolate during cold night, conduct during day	Near-zero at night
Phase change materials (PCMs)	Store heat during day, release during night	Passive
Loop heat pipes + thermal switches	Japanese and NASA collaboration	Near-zero at night
Layer 0 hibernation controller	Powers down CPU; powers only clock and survival heaters	< 1 W

THE LUNAR COMMONS is governed by CC BY-SA 4.0 (text) and CERN-OHL-S v2 (hardware). Share freely. Attribute honestly. Keep it open for all humanity.

The thermal design must handle South Pole gradients where ΔT >250°C. These techniques will be evaluated in combination during analog testing.

Radiation Hardening:

Approach	Application
Rad-hard or TMR-protected FPGAs	Layers 4.2 and 4.4
Rad-hard COTS components	Analog Devices, TI, STMicroelectronics, BAE Systems
Triple modular redundancy + voting	Critical memory cells and logic
Watchdog timers + independent power domains	Common-mode failure prevention
Full system-level radiation testing	Ground facilities

TID budget modelled for 5-year mission. SEE mitigation through component selection and error-correcting codes.

2.8 Graceful Degradation

If DAO coordination is unavailable, individual units default to independent operation: continue scheduled observation sweeps, store all signed data in local WORM storage, enter low-power hibernation between tasks, periodically attempt to re-establish connection. Upon reconnection, upload stored signed data for verification. No data is lost.

2.9 Fleet Coordination

Dimension	Protocol
Inter-unit communication	No direct peer-to-peer. All coordination mediated through DAO layer.
Coverage zone assignment	Zone definitions written into each unit's Goal ROM at manufacture. No remote reassignment.
Data aggregation	DAO ingests signed bundles, verifies signatures, and merges data.
Failure detection	Signed beacon includes power, storage, and sensor status. DAO flags absent beacons.
Collision avoidance	Proximity sensors hardwired to Layer 4.3. Speed limit 0.4 m/s.
Fleet expansion	New units: Goal ROM written, registered by DAO vote, physically activated.

THE LUNAR COMMONS is governed by CC BY-SA 4.0 (text) and CERN-OHL-S v2 (hardware). Share freely. Attribute honestly. Keep it open for all humanity.

Fleet-wide goal update	Requires physical access to Goal ROM and Emergency Council approval. Cannot be done remotely.

2.10 Scaling to Cislunar and Beyond

The same hardware-enforced observability model applies to cislunar space, asteroid resources, and Mars (same five axioms with different environmental mitigations). The core contribution — hardware-enforced, verifiable observability with no self-representation — is independent of celestial body. The Lunar Witness is the first implementation, not the last.

2.11 Live Human Observation of the Witness Fleet (Earth-Side Immersion Layers)

The Lunar Witness fleet's outputs flow outward as one-way signed broadcasts. Only timing synchronisation signals are accepted on the uplink. No command pathways reach the rover hardware.

This design enables near-live human access to observational data without compromising the Five Axioms. Observers on Earth can access full-disk Earth imagery, atmospheric spectra, lunar surface panoramas with regolith textures and horizon vistas, and outer-space views encompassing star fields and the cislunar environment.

Fundamental Latency Floor: The Earth-Moon one-way light delay of approximately 1.3 seconds sets the immutable physical minimum for any data stream.

Access Pathways (All Post-Validation and Cryptographically Signed):

1. Public open access: Signed data bundles receivable by ground stations worldwide. Decoded outputs feed public dashboards, APIs, and displays.

2. *Immersive digital extensions*: VR and AR headsets transform streams into panoramic experiences.

3. *Optional high-bandwidth neuro-perceptual layer*: [Optional — Dependent on Third-Party Technology Maturity] Ground-station decoders can map signed data into neural stimulation patterns compatible with brain-computer interfaces. As of early 2026, Neuralink's Telepathy system has enabled 21 clinical trial participants to achieve thought-based control of digital interfaces at rates matching or exceeding able-bodied mouse performance (~8-10+ bits per second), with zero reported serious device-related adverse events. The company's Blindsight initiative (FDA Breakthrough Device Designation) is progressing toward human trials for visual cortex stimulation. Complementary platforms from Paradromics and Precision Neuroscience further expand options. This layer remains strictly read-only and Earth-side.

Multi-Sensory and Shared Immersion Extensions:

· *Wearable interfaces*: Haptic suits or exoskeletons can synchronise tactile feedback with decoded streams.

· *Large-scale shared environments*: Spherical and fulldome projection theaters (e.g., Sphere

THE LUNAR COMMONS is governed by CC BY-SA 4.0 (text) and CERN-OHL-S v2 (hardware). Share freely. Attribute honestly. Keep it open for all humanity.

in Las Vegas) enable collective immersion with 360° visuals, immersive audio, haptic seating, and atmospheric effects.

Core Architectural Invariants: All immersion layers operate exclusively on Earth using publicly broadcast, cryptographically signed bundles. No additional hardware, command uplink, or internal self-representation is introduced to the Lunar Witness.

Partnership Invitation: Engagement is invited with BCI developers, haptic technology providers, exoskeleton manufacturers, planetarium operators, and accessibility organisations.

Part 3 — Limitations and Open Risks

Risk	Likelihood	Mitigation	Residual
No technical co-lead	High	Scoping study funds contractors. Sourcing via UNSW / ANU / QUT.	High
Cost escalation	Medium-High	3–5× multiplier assumed. Real missions often exceed estimates.	Medium-High
Formal verification scope	Medium	Partner with UNSW / Data61. Non-verified portions isolated via hardware.	Medium
Export control (ITAR)	Low-Medium	Phased roadmap; non-US supply chain.	Low-Medium
DAO coordination dependency	Low	Graceful degradation to local WORM storage.	Low
Physical tampering	Low-Medium	Multi-layer seals; remote imaging.	Low-Medium
Latent self emergence	Low	seL4; diverse redundancy; fault injection. Full elimination not guaranteed.	Low
Dust accumulation	Medium	EDS + seals + nanocoatings. Multi-year degradation remains a leading mode.	Medium
Thermal cycling fatigue	Medium	VCHPs + PCMs + hibernation. Component aging over 5+ years is a risk.	Medium
Radiation-induced fault	Low-Medium	TMR + rad-hard components + error correction. SEE events cannot be eliminated.	Low-Medium

THE LUNAR COMMONS is governed by CC BY-SA 4.0 (text) and CERN-OHL-S v2 (hardware). Share freely. Attribute honestly. Keep it open for all humanity.

DAO mechanical veto	Medium	5-of-7 keys. Under latency, sparse fleet, or tampering, exercising is challenging.	Medium
Immersion extensions as command pathway	Closed	Blocked by one-way broadcast, hardware validator, kill switch.	Closed

Residual risks acknowledgment: No architecture eliminates all pathways for failure, tampering, or drift. The Moon is unforgiving. Full lunar deployment requires international collaboration, Outer Space Treaty compliance, and likely government oversight.

Flight Qualification Multiplier:

Phase	Cost Estimate
Earth prototype (lab)	$280–380 k
Lunar-qualified prototype	$1.2–1.8 M
Flight unit (one-off)	$3–5 M
Fleet of 5 units	$10–15 M

Part 4 — Prior Art and Legal

Key Prior Art Claims

Claim	Description
Goal without identity	OTP Goal ROM with no self-representation.
Exploration-only rover design	No tools, arms, or excavation. Speed-limited. Proximity-interlocked.
Fully deterministic query generation	No statistical AI.
Hardware validator + kill switch stack	Four-layer independent hardware governance in series.
Separation of roles	Witness observes, DAO decides, Executor acts.
Graceful degradation	Individual units default to independent WORM storage.
Layered environmental hardening	EDS, VCHPs/PCMs, rad-hard components, seL4 pathways.
Earth-side immersion layers	One-way broadcast only. No command pathway.

THE LUNAR COMMONS is governed by CC BY-SA 4.0 (text) and CERN-OHL-S v2 (hardware). Share freely. Attribute honestly. Keep it open for all humanity.

Stress Test Summary (20 Findings)

#	Finding	Solution	Status
01	Corporate patent capture	Defensive disclosure across 4 jurisdictions	Closed
02	SaaS loophole	AGPL-3.0 + patent wall	Closed
03	DAO legal personhood	Wyoming DUNA + Swiss Stiftung	Closed
04	Premature Phase 3 declaration	4 non-waivable conditions	Closed
05	Seize the hardware	Handshake loop requires Earth pilot signal	Closed
06	Reed relay single point of failure	2-of-3 contactors + mechanical fuse	Closed
07	Peer Witness underspecification	3 mathematically distinct detection channels	Partial
08	Observatory count too low	Minimum 5 with jurisdiction diversity	Closed
09	Swiss foundation override	Subordination clause + 3-of-3 multisig	Closed
10	Global Commons Dividend capture	Non-transferable energy vouchers	Closed
11	AGPL/Linux incompatibility	System Layer Exception	Closed
12	Insurance market absence	Risk model published for insurers	Closed
13	No analog backup to DAO	100% mechanical EPC (7 keys, 5 to shutdown)	Closed
14	No generational reset	Phase 4: 75% vote every 50 years	Closed
15	No first-mover incentive	Trillion-dollar buyout + industrial sandbox	Closed
16	Electronic side-channel attack	100% mechanical keys, no electronics	Closed
17	Reflective vector beam redirection	SRV + Public Math + backscatter	Partial

THE LUNAR COMMONS is governed by CC BY-SA 4.0 (text) and CERN-OHL-S v2 (hardware). Share freely. Attribute honestly. Keep it open for all humanity.

18	EMP degradation	Passive components only, no active electronics	Closed
19	Rover repurposed for construction / weaponisation	Capability omission: no tools, arms, excavation. Speed-limited. Proximity-interlocked. Epistemic filter blocks targeting.	Closed
20	Immersion extensions as command pathways	Blocked by one-way broadcast architecture, hardware validator, mechanical kill switch, and command isolation. All extensions are strictly read-only and Earth-side after cryptographic signing and validation.	Closed

Certification status: 18 of 20 design findings fully closed. 2 findings partially closed with honest residual risks acknowledged.

Licence Summary

Scope	Licence
Hardware designs (chassis, circuits, FPGA configs)	CERN-OHL-S v2
Software components	GPL v3 with System Layer Exception
Text, governance framework, philosophical contributions	CC BY-SA 4.0

Contact

Field	Details
Name	Sandip Ghimire
X	@Vaskar017
Email	sandyghimire@gmail.com
v3.3 Archive	archive.org/details/lunar-commons-v-3-3-final
v3.3 SHA-256	4f00c4723c8faea8bf4cae90d18366f9f23af76e2d478ea81b019796fbb156c2
This Edition	archive.org/details/lunar-commons-witness-final-archive

Version History (Condensed)

Version	Date	Summary of Changes
v3.3	1 April 2026	Selfless AI, Peer Witness, DUNA DAO, physical logic gates

THE LUNAR COMMONS is governed by CC BY-SA 4.0 (text) and CERN-OHL-S v2 (hardware). Share freely. Attribute honestly. Keep it open for all humanity.

v4.0	4 April 2026	Goal without identity. Fully deterministic. Exploration-only. (not published)
v5.3	9 April 2026	Layered environmental hardening. Residual risk quantification. Earth-side immersion layers with BCI specificity and cislunar vistas. Full stress-test table restored. Research-frontier labels. Archive final.

THE LUNAR COMMONS is governed by CC BY-SA 4.0 (text) and CERN-OHL-S v2 (hardware). Share freely. Attribute honestly. Keep it open for all humanity.